

Description

METHOD AND STRUCTURE FOR IMPROVED MOSFETs USING POLY/SILICIDE GATE HEIGHT CONTROL

BACKGROUND OF INVENTION

[0001] The invention generally relates to methods for manufacturing a semiconductor device with improved device performance, and more particularly to methods for manufacturing semiconductor devices which impose tensile and compressive stresses in the substrate of the device during device fabrication.

[0002] Mechanical stresses within a semiconductor device substrate can modulate device performance. That is, stresses within a semiconductor device are known to enhance semiconductor device characteristics. Thus, to improve the characteristics of a semiconductor device, tensile and/or compressive stresses are created in the channel of the n-type devices (e.g., NFETs) and/or p-type devices (e.g., PFETs). However, the same stress component, either ten-

sile stress or compressive stress, discriminatively affects the characteristics of an n-type device and a p-type device.

[0003] In order to maximize the performance of both NFETs and PFETs within integrated circuit (IC) chips, the stress components should be engineered and applied differently for NFETs and PFETs. That is, because the type of stress which is beneficial for the performance of an NFET is generally disadvantageous for the performance of the PFET. More particularly, when a device is in tension (in the direction of current flow in a planar device), the performance characteristics of the NFET are enhanced while the performance characteristics of the PFET are diminished. To selectively create tensile stress in an NFET and compressive stress in a PFET, distinctive processes and different combinations of materials are used.

[0004] For example, a trench isolation structure has been proposed for forming the appropriate stresses in the NFETs and PFETs, respectively. When this method is used, the isolation region for the NFET device contain a first isolation material which applies a first type of mechanical stress on the NFET device in a longitudinal direction (parallel to the direction of current flow) and in a trans-

verse direction (perpendicular to the direction of current flow). Further, a first isolation region and a second isolation region are provided for the PFET and each of the isolation regions of the PFET device applies a unique mechanical stress on the PFET device in the transverse and longitudinal directions.

[0005] Alternatively, liners on gate sidewalls have been proposed to selectively induce the appropriate strain in the channels of the FET devices (see Ootsuka et al., IEDM 2000, p.575, for example). By providing liners the appropriate stress is applied closer to the device than the stress applied as a result of the trench isolation fill technique.

[0006] While these methods do provide structures that have tensile stresses being applied to the NFET device and compressive stresses being applied along the longitudinal direction of the PFET device, they may require additional materials and/or more complex processing, and thus, resulting in higher cost. Further, the level of stress that can be applied in these situations is typically moderate (i.e., on the order of 100s of MPa). Thus, it is desired to provide more cost-effective and simplified methods for creating large tensile and compressive stresses in the channels NFETs and PFETs, respectively.

SUMMARY OF INVENTION

[0007] In a first aspect, this invention provides a method for manufacturing an integrated circuit comprising a plurality of semiconductor devices including an n-type transistor and a p-type transistor. The method involves depositing oxide fill on the n-type transistor and the p-type transistor and chemical/mechanical polishing the deposited oxide fill such that a gate stack of the n-type transistor and a gate stack of the p-type transistor, the n-type transistor and the p-type transistor having spacers are surrounded with oxide. The method further involves etching a portion of the polysilicon from a gate of the p-type transistor, depositing a low resistance material (e.g., Co, Ni, Ti, or other similar metals) on the n-type transistor and the p-type transistor, and heating the integrated circuit such that the deposited material reacts with the polysilicon of the n-type transistor and the polysilicon of the p-type transistor to form silicide. The silicide formed on the p-type polysilicon imposes compressive mechanical stresses along the longitudinal direction of the channel of the p-type transistor.

[0008] In a second aspect, this invention separately provides an integrated circuit, comprising a p-type transistor having a

polysilicon layer and a material with a low-resistance thereon, an n-type transistor having a polysilicon layer and a material with a low-resistance thereon. A height of the polysilicon layer of the p-type transistor is less than a height of a polysilicon layer of the n-type transistor.

[0009] In another aspect, this invention separately provides a method for manufacturing an integrated circuit comprising a plurality of semiconductor devices including an n-type transistor and a p-type transistor. The method comprising forming a polysilicon layer on the n-type transistor and the p-type transistor, wherein the polysilicon layer on the p-type transistor has a shorter height than the polysilicon layer on the n-type transistor. A low-resistance material is deposited on the p-type transistor and the n-type transistor. The integrated circuit is heated such that the deposited low-resistance material reacts with the polysilicon of the p-type transistor and the polysilicon of the n-type transistor. In the methods, compressive mechanical stresses are formed along a longitudinal direction of a channel of the p-type transistor.

BRIEF DESCRIPTION OF DRAWINGS

[0010] Figure 1 depicts desired stress states for PFETs and NFETs;

- [0011] Figures 2(a) through 2(h) depict an exemplary process for forming the MOSFETs according to the invention;
- [0012] Figure 3 depicts stresses in a gate structure with a full poly height according to the invention; and
- [0013] Figure 4 depicts stresses in a gate structure with an etched poly stack according to the invention.

DETAILED DESCRIPTION

- [0014] The invention provides a method for fabricating devices with improved performance characteristics. In one aspect of the invention, polysilicon is over-etched approximately half way down a gate stack structure of a PFET using a poly etch technique before silicidation of the PFET. The silicidation process of the over-etched polysilicon places the critical part of the silicon region under the gate oxide under compression. Thus, by over-etching the polysilicon of the PFET, according to the invention, the bending of the isolation is minimized and a compressive stress state is more directly relayed to the channel region. Generally, the invention uses over-etched polysilicon gates to create compressive stresses in the channels of the PFETs. By providing tensile stresses to the channel of the NFET and compressive stresses to the channel of the PFET the charge mobility and drive current along the channels of

each device is enhanced.

[0015] By implementing the stresses using a polysilicon over-etch, this invention provides the implementation of stresses which are self-aligned to the gate, whereas in the isolation induced stress structures, the stresses are not self-aligned to the gate. This invention also provides stress levels in the silicon under the gate which are much larger than the isolation-based or liner-based approaches.

[0016] In this invention, oxide fill is used to provide confinement to a gate stack. When the gates are confined laterally using oxide fill, the expansion of the silicide in the gate stack is relatively confined. Further, when a gate that is laterally surrounded with oxide fill is subjected to silicidation, the gate channel is put under tension. In particular, the gate channel is put under tension because silicide reacted thereon expands outwardly at the top. This results in compression in the top part of the isolation. At the same time, by inducing bending stresses at the bottom part of the isolation, next to the polysilicon under tension, the channel receives large tensile stresses. These stress levels are relatively high (on the order of 1–2GPa) and these tensile stresses are beneficial to the NFET drive cur-

rents.

[0017] However, these tensile stresses significantly hinder the PFET drive currents. In this invention, the stresses in the PFET are reduced and/or made to be compressive stresses rather than tensile stresses in order to improve the performance of the PFET. Thus, this invention provides a method for providing longitudinal compressive stresses along the channel of the PFET while providing tensile stresses along the channel of the NFET to improve the performance of the devices.

[0018] Figure 1 illustrates desired stress states for improving the performance of PFETs and NFETs (see Wang et al., IEEE Tran. Electron Dev., v.50, p.529, 2003). In Figure 1, an NFET and a PFET are shown to have a source region, a gate region and a drain region. The NFET and PFET are shown to have arrows extending outward from the active area to illustrate tensile stresses. The arrows extending inward toward the PFET device are illustrative of compressive forces. More specifically, the outwardly extending arrows, shown extending from the NFET, illustrate a tensile stress that is desired in the transverse and longitudinal directions of the device. Similarly, the inwardly extending arrows, shown with relation to the PFET, illustrate a de-

sired longitudinal compressive stress. The range of stresses needed to influence device drive currents is of the order of a few hundred MPa to a few GPa. The width and the length of the active area of each device are represented by "W" and "L", respectively. It should be understood that each of the longitudinal or transverse stress components could be individually tailored to provide the performance enhancements for both devices (i.e., the NFET and the PFET).

[0019] Figures 2(a) through 2(h) depict a general exemplary process for forming the MOSFETs according to this invention. Figures 2(a) through 2(d) explain processes that are known and thus, any known applicable process may be used. Figure 2(a) illustrates the structure after the shallow trenches 5 (STI) are formed. An SOI (silicon-on-insulator) wafer which has a stack of silicon 1, buried oxide 2, and a silicon layer 3 is used. Generally, to form the STI on SOI wafers, a thin ($\sim 50\text{\AA}$) layer of silicon dioxide SiO_2 (pad oxide) (not shown) is grown on the silicon layer 3, which is on the buried oxide layer 2 on the silicon substrate 1, by reacting silicon and oxygen at high temperatures. A thin layer (about 1000\AA to about 2500\AA) of pad silicon nitride (Si_3N_4) (not shown) is then deposited using chemical vapor

deposition (CVD). Next, the patterned photoresist with a thickness of about 0.1 microns to about 1.0 microns is deposited, and the structure is exposed and developed to define the trench areas 5. Next, the exposed SiO_2 and the Si_3N_4 are etched using reactive ion etching (RIE). Next, an oxygen plasma is used to burn off the photoresist layer. A wet etch is used to remove the pad Si_3N_4 and pad oxide. Then an oxide layer is deposited to fill the trenches and the surface oxide is removed using chemical mechanical polishing (CMP). This completes the formation of STI as seen in Figure 2(a).

[0020] Next, a sacrificial oxide (not shown) of about 50Å is grown on the silicon. Then, as shown in Figure 2(b) an n-well 10 and a p-well 12 are formed. Patterned photoresist layers may be used to successively form the n-well 10 (using, for example, multiple implants of Phosphorous ions) and the p-well 12 (using, for example, multiple implants of Boron ions). The well implants are then optionally annealed. The sacrificial oxide layer is then removed using a wet HF solution, such that a clean silicon surface is left behind.

[0021] Next, as shown in Figure 2(b), a gate oxide layer 14 of about 10Å to about 100Å is grown. On the gate oxide layer

14, a polysilicon layer 16 is deposited using CVD to a thickness of about 500Å to about 1500Å to form the gate electrodes 18 and 20 shown in Figure 2(c). Patterned photoresist layers (not shown) are used to define the gate electrodes. RIE is used to etch the exposed portions of the polysilicon layer 16 and the photoresist patterns are stripped away in order to complete formation of the gate stack of the n-type transistor 17 and the gate stack of the p-type transistor 19.

[0022] Figure 2(c) shows the formed gate electrodes 18 and 20. A thin layer of oxide 15 is then grown on the remaining polysilicon. Patterned photoresist layers (not shown), which are later removed, are used to successively tip (and halo countering doping implants) implant the n-type and p-type transistors. For n-type transistors, a very shallow and low dose implant of arsenic ions, for example, may be used to form the p-tip 22 (while a Boron implant, for example, may be used for halos). For p-type transistors, a very shallow and low dose implant of BF_2 ions, for example, may be used to form n-tip 24 (while an arsenic implant may, for example, be used for halos).

[0023] Next, as shown in Figure 2(c), spacers 26 are formed by depositing a silicon nitride layer (not shown) using CVD to

a thickness of about 100Å to about 1000Å and then etching the nitride from the regions other than the sidewalls of the gate. Patterned photoresist layers (not shown), which are removed prior to the next stage of the process, are used to successively create the source/drain regions of the transistors. In Figure 2(d), for the n-type transistors, a shallow and high-dose of arsenic ions, for example, may be used to form the source/drain regions 28 while the p-type transistors are covered with the corresponding photoresist layer. For the p-type transistors, a shallow and high dose of BF_2 ions, for example, may be used to form the source/drain regions 30 while the n-type transistors are covered with the corresponding photoresist layer. An anneal is then used to activate the implants. The exposed oxide on the structure is then stripped by dipping the structure in HF in order to expose bare silicon in the source, gate and drain regions of the transistors.

[0024] Still referring to Figure 2(d), metal 32 is deposited to a thickness of about 30Å to about 200Å across the wafer surface in order to form silicide. The silicide could be formed from reacting the underlying with any deposited metal such as Co, Hf, Mo, Ni, Pd, Pt, Ta, Ti, W, and Zr. In the regions, such as, the source, drain and gate regions,

where the deposited metal is in contact with silicon, the deposited metal reacts with the silicon to form silicide. In the other regions (i.e., where the deposited metal is not in contact with silicon), the deposited metal remains unchanged. This process aligns the silicide to the exposed silicon and is called "self-aligned silicide" or salicide. The unreacted metal is then removed using a wet etch while the formed silicide 34 remains, as shown in Figure 2(e).

[0025] In methods according to the invention, as shown in Figure 2(e), an oxide fill 36 followed by chemical mechanical polishing is used to planarize the surface. Next, as shown in Figure 2(f), the silicide 34 on top of the polysilicon is removed using a selective etch. Then, as shown in Figure 2(g), the n-type device of the structure is covered with a mask 38. The mask may be, for example, photoresist material. Next, the polysilicon of the p-type device is partially etched about half-way down the gate stack. The ratio of the height of the polysilicon of the n-type device to the height of the polysilicon of the p-type device may be 2:1. More particularly, gate height stacks are generally between about 500Å to about 1500Å. In this case, the p-type gate polysilicon is etched down to about 250Å to about 750Å. Standard wet or dry etch known in the art for

polysilicon or doped polysilicon may be used. As shown in Figure 2(h) after partially etching the polysilicon of the p-type device, the height of the polysilicon of the p-type device is shorter than the height of the polysilicon of the n-type device.

[0026] After removing the mask 38 over the n-type device, a metal or low resistance material 40, such as, for example, Co, HF, Mo, Ni, Pd₂, Pt, Ta, Ti, W, and Zr is deposited over the wafer via, for example, evaporation, sputtering, or CVD techniques. In particular, Ni, for example, may be deposited over the polysilicon of the n-type and p-type devices and reacted to form silicide at different heights of the stack for the n-type and p-type devices, respectively.

[0027] Next, the structure is heated to temperature of about 300°C to about 1000°C to allow the deposited silicide material to react with the polysilicon. During sintering, silicide only forms in the regions where metal is in direct contact with silicon or polysilicon. After the reaction, the silicide thickness can be anywhere from about 50 Å to about 500 Å, depending on the initial metal thickness and temperature and time of the anneal. The remaining unreacted silicide forming material is then removed, for example, with a selective etch without damaging the formed

silicide. The remaining formed silicide is automatically self-aligned with the gate of the MOS device. As shown in Figure 2(g), the height of the silicide material on the p-type transistor is greater than a height of the silicide material on the n-type transistor. After this, the standard middle of the line (e.g., passivation and contact formation) and back end of the line (various interconnect metal, via, and interlevel dielectrics) processes are performed.

[0028] In another embodiment of the invention, silicide forming material may be deposited over the gate 18 of the n-type device and the gate 20 of the p-type device. For example, silicide forming material may be deposited to a height of about 70Å to about 200Å. Next, chemical/mechanical polishing may be performed to polish into a planar surface the formed silicide. It should be understood by those of skill in the art that it is also possible to use co-evaporation, co-sputtering, or sputtering to simultaneously deposit both silicon and metal onto the polysilicon surface. After this, the standard middle of the line (e.g., passivation and contact formation) and back end of the line (various interconnect metal, via, and interlevel dielectrics) processes are performed.

[0029] It should be understood that this invention is readily ap-

plicable to bulk or layered SiGe substrates. It should also be understood that this invention may also be used with damascene gate structures, which have been proposed for use of high k dielectric gate oxides.

[0030] Figure 3 shows the stresses in the gate structure when full poly height is used for the PFET. The dashed lines represent tensile stress and the solid lines represent compressive stress. As can be seen from Figure 3, with full poly stack height, in the channel area of the PFET, only tensile stress is shown.

[0031] Figure 4 shows the stresses in the gate structure when an etched poly stack is used for the PFET in accordance with this invention. Similar to Figure 3, the dashed lines represent the tensile stress and the solid lines represent the compressive stress. As can be seen from Figure 4, with an etched poly stack for the PFET, in the channel area below the gate, there is primarily compressive stress. As discussed above with regards to Figure 1, in PFETs, a longitudinal compressive stress is desired. The typical range for the desired compressive/tensile stresses is on the order of a few hundred MPa to a few GPa. The invention can produce very high compressive stresses and tensile stresses in the channels of the PFET and NFET devices, re-

spectively.

[0032] By providing tensile stresses to the channel of the NFET and compressive stresses to the channel of the PFET the charge mobility along the channels of each device is enhanced. Thus, as described above, the invention provides a method for providing compressive stresses along the longitudinal direction of the channel by over-etching the polysilicon of the PFET prior to silicidation.

[0033] While the invention has been described in terms of embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the appended claims.